PRESERVERS FOR NORMS OF LIE PRODUCT

CHI-KWONG LI, EDWARD POON, AND NUNG-SING SZE

ABSTRACT. Let $\|\cdot\|$ be a unitary similarity invariant norm on the set M_n of $n \times n$ complex matrices. A description is obtained for surjective maps ϕ on M_n satisfying $\|AB - BA\| = \|\phi(A)\phi(B) - \phi(B)\phi(A)\|$ for all $A, B \in M_n$. The general theorem covers the special cases when the norm is one of the Schatten p-norms, the Ky-Fan k-norms, or the k-numerical radii.

1. Introduction

Let M_n be the set of $n \times n$ matrices. A norm $\|\cdot\|$ is called a *unitary similarity invariant norm* if $\|UAU^*\| = \|A\|$ for all $A \in M_n$ and unitary $U \in M_n$, and is called a *unitarily invariant norm* if $\|UAV\| = \|A\|$ for all $A \in M_n$ and unitary $U, V \in M_n$. Clearly, any unitarily invariant norm is a unitary similarity invariant norm. To understand a normed vector space, researchers study maps preserving the norms. Linear maps ϕ satisfying $\|\phi(A)\| = \|A\|$ are known as *linear isometries* for the norm; maps ϕ satisfying $\|\phi(A) - \phi(B)\| = \|A - B\|$ are known as *distance preserving maps.* Linear isometries for unitarily invariant norms and unitary similarity invariant norms are quite well studied; see [1, 5] and the references therein. For instance, linear isometries for unitarily invariant norms not equal to multiples of the Frobenius norm on M_n always have the form

$$A \mapsto UAV \quad \text{or} \quad A \mapsto UA^tV$$

for some unitary $U, V \in M_n$.

Since M_n is an algebra, researchers also study multiplicative maps ϕ satisfying $\|\phi(A)\| = \|A\|$, or maps on a subset of M_n satisfying $\|\phi(A)\phi(B)\| = \|AB\|$; see [3, 2]. For instance, it was shown in [2] that for a unitary similarity invariant norm $\|\cdot\|$ on M_n and a subset $\mathcal{S}$ of M_n containing all rank one idempotents, if a map $\phi : \mathcal{S} \to M_n$ satisfies $\|\phi(A)\phi(B)\| = \|AB\|$ for all $A, B \in \mathcal{S}$, then ϕ has the form

$$A \mapsto \mu_A UAU^* \quad \text{or} \quad A \mapsto \mu_A U\overline{A}U^* \tag{1.1}$$

for some unitary U and complex unit μ_A, depending on A.

In this paper, we determine the structure of surjective maps $\phi : M_n \to M_n$ such that for any $A, B \in M_n$,

$$\|[\phi(A), \phi(B)]\| = \|[A, B]\|, \tag{1.2}$$

2000 *Mathematics Subject Classification.* Primary 15A60, 46B04.
Key words and phrases. Lie product, unitarily invariant and unitary similarity invariant norms.

where $[A, B] = AB - BA$ is the Lie product of A and B. In Section 2, we show that if $\phi : M_n \to M_n$ is a surjective map satisfying (1.2), then there is a unitary $U \in M_n$ and a subset $\mathcal{N}_n$ of normal matrices in M_n such that ϕ has the form

$$\phi(A) = \begin{cases} \mu_A U A^\dagger U^* + \nu_A I_n & A \in M_n \setminus \mathcal{N}_n, \\ \mu_A U (A^\dagger)^* U^* + \nu_A I_n & A \in \mathcal{N}_n, \end{cases}$$

where $\mu_A, \nu_A \in \mathbb{C}$ with $|\mu_A| = 1$, depending on A, and $A \mapsto A^\dagger$ denotes one of the following maps: $A \mapsto A$, $A \mapsto \overline{A}$, $A \mapsto A^t$ or $A \mapsto A^*$. The set $\mathcal{N}_n$ depends on the given norm $\|\cdot\|$. For unitarily invariant norms, we characterize $\mathcal{N}_n$ in terms of the norm $\|\cdot\|$ in section 3. In particular, if $\|\cdot\|$ is the Frobenius norm $\|A\|_F = (\operatorname{tr} A^*A)^{1/2}$, then $\mathcal{N}_n$ can be any subset of the set of normal matrices. The situation is more intricate for unitary similarity invariant norms which are not unitarily invariant. In Section 4, we consider a class of norms of this nature, namely, the k-numerical radius for $k \in \{1, \ldots, n-1\}$. In such cases, we show that $\mathcal{N}_n$ is always empty.

2. Unitary similarity invariant norm

In this section, let $\|\cdot\|$ be a unitary similarity invariant norm on the set M_n of $n \times n$ complex matrices. Our main result is the following.

Theorem 2.1. *Suppose $n \geq 3$, and $\phi : M_n \to M_n$ is a surjective map satisfying*

$$\|[\phi(A), \phi(B)]\| = \|[A, B]\|.$$

Then there is a unitary matrix U and a subset $\mathcal{N}_n$ of normal matrices such that ϕ has the form

$$\phi(A) = \begin{cases} \mu_A U A^\dagger U^* + \nu_A I_n & A \in M_n \setminus \mathcal{N}_n, \\ \mu_A U (A^\dagger)^* U^* + \nu_A I_n & A \in \mathcal{N}_n, \end{cases}$$

where $\mu_A, \nu_A \in \mathbb{C}$ with $|\mu_A| = 1$, depending on A. Here, $A^\dagger = A$, $\overline{A}$, A^t or A^*.

We need the following result from Šemrl [7] to prove the above theorem.

Theorem 2.2. *Suppose $n \geq 3$, and $\phi : M_n \to M_n$ is a bijective map satisfying*

$$[A, B] = 0_n \quad \iff \quad [\phi(A), \phi(B)] = 0_n.$$

Let Γ be the set of matrices A such that the Jordan form of A only has Jordan blocks of sizes 1 or 2. Then there are an invertible matrix S, an automorphism σ of the complex field and a regular locally polynomial map $A \mapsto p_A(A)$ such that

$$\phi(A) = S(p_A(A_\sigma^\dagger))S^{-1} \quad \textit{for all } A \in \Gamma. \tag{2.1}$$

Here, X_σ is the matrix whose (i, j)-entry is $\sigma(X_{ij})$, and $A^\dagger = A$ or A^t.

Denote by $\sigma(A)$ the spectrum of A and $N(A)$ the null space of A.

Lemma 2.3. *For any two matrices A and B, if*

$$\|[A,X]\| = \|[B,X]\| \quad \text{for all rank one } X \in M_n, \tag{2.2}$$

then there are $\mu,\nu \in \mathbb{C}$ with $|\mu| = 1$ such that one of the following holds with $\hat{A} = \mu A + \nu I_n$.

(a) *$\sigma(B) = \sigma(\hat{A})$ and for any $\lambda \in \sigma(\hat{A})$,*

$$N(B - \lambda I_n) = N(\hat{A} - \lambda I_n) \quad \text{and} \quad N(B^t - \lambda I_n) = N(\hat{A}^t - \lambda I_n).$$

(b) *The eigenvalues of A are not collinear, $\sigma(B) = \overline{\sigma(\hat{A})}$ and for any $\lambda \in \sigma(\hat{A})$,*

$$N(B - \overline{\lambda} I_n) = N(\hat{A} - \lambda I_n) \quad \text{and} \quad N(B^t - \overline{\lambda} I_n) = N(\hat{A}^t - \lambda I_n).$$

Proof. Note that for any rank one matrix $X = xy^t$, $[C,X] = 0$ if and only if x and y^t are the right and left eigenvectors of C corresponding to the same eigenvalue. To see this, as $[C,X] = (Cx)y^t - x(y^tC)$, then $[C,X] = 0$ if and only if $Cx = \lambda x$ and $y^tC = \lambda y^t$ for some $\lambda \in \mathbb{C}$.

Suppose A and B satisfy (2.2). By the above observation on rank one matrices, A and B must have the same sets of left and right eigenvectors. Furthermore, x_1 and x_2 are the right eigenvectors of A corresponding to the same eigenvalue if and only if the two eigenvectors correspond to the same eigenvalue of B. Thus, the eigenvalues of A and B have the same geometric multiplicity.

Let $\lambda_1,\dots,\lambda_k$ be the distinct eigenvalues of A with $x_1,\dots,x_k$ and $y_1,\dots,y_k$ being the right and left eigenvectors. Also for each pair of eigenvectors x_i and y_i^t, let γ_i be the corresponding eigenvalue of B. Take $X_{ij} = x_i y_j^t$. Then $AX_{ij} = \lambda_i X_{ij}$ and $X_{ij}A = \lambda_j X_{ij}$. So for any $1 \le i,j \le k$,

$$\|[A,X_{ij}]\| = \|\lambda_i X_{ij} - \lambda_j X_{ij}\| = |\lambda_i - \lambda_j| \|X_{ij}\|.$$

Similarly, $\|[B,X_{ij}]\| = |\gamma_i - \gamma_j| \|X_{ij}\|$. Therefore, $|\lambda_i - \lambda_j| = |\gamma_i - \gamma_j|$ for all $1 \le i,j \le k$. Then there are $\mu,\nu \in \mathbb{C}$ with $|\mu| = 1$ such that either

(1) $\gamma_i = \mu\lambda_i + \nu$ for all $1 \le i \le k$; or
(2) the eigenvalues of A are non-collinear and $\overline{\gamma}_i = \mu\lambda_i + \nu$ for all $1 \le i \le k$.

Then the result follows with $\hat{A} = \mu A + \nu I_n$. □

Lemma 2.4. *Suppose A and B commute and satisfy (2.2). If A has at least two distinct eigenvalues, then there are $\mu,\nu \in \mathbb{C}$ with $|\mu| = 1$ such that*

(a) *$B = \mu A + \nu I_n$, or*
(b) *A is normal with non-collinear eigenvalues and $B = \mu A^* + \nu I_n$.*

Proof. As A and B commute, there is a unitary matrix U such that both U^*AU and U^*BU are upper triangular, see [6, Theorem 2.3.3]. By replacing (A,B) with (U^*AU, U^*BU), we may assume that A and B are upper triangular.

As A and B satisfy (2.2), Lemma 2.3 holds. Suppose Lemma 2.3(a) holds with $\hat{A} = \mu A + \nu I_n$. Notice that $\sigma(B) = \sigma(\hat{A})$ and

$$\|[\hat{A}, X]\| = \|[\mu A + \nu I_n, X]\| = \|[B, X]\| \quad \text{for all rank one } X \in M_n.$$

Suppose λ is an eigenvalue of $\hat{A}$ and $y \in N(\hat{A}^t - \lambda I_n)$. For any $z \in \mathbb{C}^n$, let $Z = zy^t$. Then $Z\hat{A} = \lambda Z$ and $[\hat{A}, Z] = (\hat{A} - \lambda I_n)Z$. Note that $(\hat{A} - \lambda I_n)Z$ has rank at most one and $\operatorname{tr}((\hat{A} - \lambda I_n)Z) = \operatorname{tr}([\hat{A}, Z]) = 0$, so $(\hat{A} - \lambda I_n)Z$ is unitarily similar to $\|(\hat{A} - \lambda I_n)z\|\|y^t\|E_{12}$, where $\|y^t\| = \|y\|$ is the ℓ_2-norm of the vector y. Thus,

$$\|[\hat{A}, Z]\| = \|(\hat{A} - \lambda I_n)z\|\,\|y^t\|\|E_{12}\|.$$

Similarly, $\|[B, Z]\| = \|(B - \lambda I_n)z\|\,\|y^t\|\|E_{12}\|$. Hence,

$$\|(\hat{A} - \lambda I_n)z\| = \|(B - \lambda I_n)z\| \quad \text{for all } z \in \mathbb{C}^n \text{ and } \lambda \in \sigma(\hat{A}).$$

Now as

$$\begin{aligned} & z^*\hat{A}^*\hat{A}z - 2\mathrm{Re}\,(\overline{\lambda} z^*\hat{A}z) + |\lambda|^2 z^*z = \|(\hat{A} - \lambda I_n)z\|^2 \\ = \quad & \|(B - \lambda I_n)z\|^2 = z^*B^*Bz - 2\mathrm{Re}\,(\overline{\lambda} z^*Bz) + |\lambda|^2 z^*z, \end{aligned}$$

this implies that

$$2\mathrm{Re}\,(\overline{\lambda} z^*(\hat{A} - B)z) = z^*(\hat{A}^*\hat{A} - B^*B)z \quad \text{for all } z \in \mathbb{C}^n \text{ and } \lambda \in \sigma(\hat{A}).$$

As A has at least two distinct eigenvalues, so does $\hat{A}$. Taking any $\lambda, \gamma \in \sigma(\hat{A})$ with $\lambda \neq \gamma$, we have

$$2\mathrm{Re}\,(\overline{\lambda} z^*(\hat{A} - B)z) = z^*(\hat{A}^*\hat{A} - B^*B)z = 2\mathrm{Re}\,(\overline{\gamma} z^*(\hat{A} - B)z).$$

Thus, $W((\overline{\lambda - \gamma})(\hat{A} - B)) \subseteq i\mathbb{R}$, where $W(X)$ is the numerical range of X. Then $(\overline{\lambda - \gamma})(\hat{A} - B)$ is a skew-Hermitian matrix and hence $\hat{A} - B$ is a diagonal matrix. Now for any $1 \le i \le n$, $b_{ii} \in \sigma(B) = \sigma(\hat{A})$. Also the ith entry of $(B - b_{ii}I_n)e_i$ is zero while only the ith entry of $(\hat{A} - B)e_i$ can be nonzero. Then

$$\begin{aligned} \|(B - b_{ii}I_n)e_i\|^2 &= \|(\hat{A} - b_{ii}I_n)e_i\|^2 \\ &= \|(B - b_{ii}I_n)e_i + (\hat{A} - B)e_i\|^2 \\ &= \|(B - b_{ii}I_n)e_i\|^2 + \|(\hat{A} - B)e_i\|^2. \end{aligned}$$

Thus, $(\hat{A} - B)e_i = 0$ for all $1 \le i \le n$ and hence $B = \hat{A}$.

Now suppose Lemma 2.3(b) holds. Then by a similar argument, we can show that

$$\|(\hat{A} - \lambda I_n)z\| = \|(B - \overline{\lambda} I_n)z\| \quad \text{for all } \lambda \in \sigma(\hat{A}) \text{ and } z \in \mathbb{C}^n \tag{2.3}$$

and so $(\overline{\lambda - \gamma})\hat{A} - (\lambda - \gamma)B$ is a skew-Hermitian matrix. Then

$$(\overline{\lambda - \gamma})T_A - (\lambda - \gamma)T_B = 0,$$

or equivalently, $T_B = \frac{\overline{\lambda-\gamma}}{\lambda-\gamma} T_A$, where T_A and T_B are the strictly upper triangular parts of A and B. Now as the eigenvalues of A and hence $\hat{A}$ are not collinear, we can always find another $\omega \in \sigma(\hat{A})$ such that $\frac{\overline{\lambda-\omega}}{\lambda-\omega} \neq \frac{\overline{\lambda-\gamma}}{\lambda-\gamma}$. Then the above equation is possible only if $T_A = T_B = 0$. In this case, A and B are both diagonal and hence normal. Also (2.3) implies that $\hat{A} = \overline{B}$. □

From Lemma 2.4, we have the following consequence for diagonalizable matrices.

Corollary 2.5. *Suppose A and B satisfy (2.2) and A is diagonalizable. Then there are $\mu, \nu \in \mathbb{C}$ with $|\mu| = 1$ such that*

(a) *$B = \mu A + \nu I_n$, or*
(b) *A is normal with non-collinear eigenvalues and $B = \mu A^* + \nu I_n$.*

Proof. Suppose A is diagonalizable. Then $A = SDS^{-1}$ for some invertible S and diagonal D. By Lemma 2.3, $B = S(\mu D + \nu I_n)S^{-1}$ or $B = S(\mu \overline{D} + \nu I_n)S^{-1}$. If A has only one eigenvalue, then A is a scalar matrix and so is B. Then the result follows. Suppose A has at least two eigenvalues. As A and B commute, the result now follows by Lemma 2.4. □

Lemma 2.6. *For any two matrices A and B, if*

$$\|[A, X]\| = \|[B, X]\| \quad \text{for all } X \in M_n, \tag{2.4}$$

then there are $\mu, \nu \in \mathbb{C}$ with $|\mu| = 1$ such that

(a) *$B = \mu A + \nu I_n$, or*
(b) *A is normal with non-collinear eigenvalues and $B = \mu A^* + \nu I_n$.*

Proof. Suppose A and B satisfy (2.4). Then clearly A and B commute. If A has at least two eigenvalues, then the result follows from Lemma 2.4.

Suppose A has only one eigenvalue, say λ. Then by Lemma 2.3, B has one eigenvalue only, say γ. Write $A = SJS^{-1} + \lambda I_n$, where S is invertible and $J = J_{n_1} \oplus \cdots \oplus J_{n_s}$ is the Jordan block form of A with $n_1 \geq \cdots \geq n_s$. Now as A and B satisfy (2.4), A and B have the same set of commuting matrices. Then $B = Sp(J)S^{-1} + \gamma I_n$ for some polynomial p of degree at most $m = n_1 - 1$ with $p(0) = 0$.

By a similar argument as in Lemma 2.4, we can show that

$$\|(B - \gamma I_n)z\| = \|(A - \lambda I_n)z\| \quad \text{for all } z \in \mathbb{C}^n.$$

Then there is a unitary matrix W such that

$$Sp(J)S^{-1} = (B - \gamma I_n) = W(A - \lambda I_n) = WSJS^{-1}.$$

Write $S = UT$ for unitary U and upper triangular T, $V = U^*WU$ and $p(x) = \sum_{i=1}^{m} c_i x^i$. Then we have

$$Tp(J)T^{-1} = VTJT^{-1}. \tag{2.5}$$

Notice that both $Tp(J)T^{-1}$ and TJT^{-1} are strictly upper triangular. Furthermore, the first $n_1 - 1$ entries in the super-diagonal of $Tp(J)T^{-1}$ are c_1 times the corresponding $n_1 - 1$ super-diagonal entries of TJT^{-1}.

As V is unitary, we must have $|c_1| = 1$ and $V = c_1 I_{n_1-1} \oplus V_1$ for some unitary $V_1 \in M_{n-n_1+1}$. Now comparing the leading $n_1 \times n_1$ principal submatrices in (2.5), we have

$$T_1 p(J_{n_1}) T_1^{-1} = (c_1 I_{n_1-1} \oplus [v_{n_1,n_1}]) T_1 J_{n_1} T_1^{-1} = c_1 T_1 J_{n_1} T_1^{-1},$$

where T_1 is the leading $n_1 \times n_1$ principal submatrix of T. Thus, $T_1 \left(\sum_{i=2}^m c_i J_{n_1}^i\right) T_1^{-1} = 0$ and so $\sum_{i=2}^m c_i J_{n_1}^i = 0$. Hence, $c_2 = \cdots = c_m = 0$. Then $p(x) = c_1 x$ and so $B = c_1 A + (\gamma - c_1\lambda) I_n$. $\square$

We are now ready to present the following.

Proof of Theorem 2.1.

First we assume that ϕ is bijective. Since

$$\|[A, B]\| = \|[\phi(A), \phi(B)]\| \quad \text{for all } A, B \in M_n,$$

by Theorem 2.2, ϕ has the form (2.1) with $A^\dagger = A$ or A^t. In particular, for any rank one matrix $R \in M_n$, there are $\mu_R, \nu_R \in \mathbb{C}$ with $\mu_R \neq 0$ such that

$$\phi(R) = S(\mu_R R_\sigma^\dagger + \nu_R I_n) S^{-1}.$$

Without loss of generality, we may assume that $\mu_R > 0$ and $\nu_R = 0$.

We here consider only the case when $A^\dagger = A$. The case when $A^\dagger = A^t$ is similar. Fix an orthonormal basis $\{x_1, \ldots, x_n\}$ and define $X_{ij} = x_i x_j^*$. Take $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{C}^n$ and let $A = \sum_{j=1}^n \alpha_j X_{j1}$. For $k = 2, \ldots, n$,

$$\|\mu_A \mu_{X_{kk}} \sigma(\alpha_k) S(X_{k1})_\sigma S^{-1}\| = \|[\phi(A), \phi(X_{kk})]\| = \|[A, X_{kk}]\| = \|\alpha_k X_{k1}\|.$$

In particular, if $\alpha_2 \neq 0$ and $\alpha_3 = 1$, we can deduce that

$$\left|\frac{\alpha_2}{\sigma(\alpha_2)}\right| \cdot \frac{\|X_{21}\|}{\|\mu_{X_{22}} S(X_{21})_\sigma S^{-1}\|} = \mu_A = \left|\frac{1}{\sigma(1)}\right| \cdot \frac{\|X_{31}\|}{\|\mu_{X_{33}} S(X_{31})_\sigma S^{-1}\|}. \tag{2.6}$$

Thus, $\left|\frac{\alpha_2}{\sigma(\alpha_2)}\right|$ is a constant. Since σ is an automorphism on $\mathbb{C}$, it is either the identity map $\lambda \mapsto \lambda$ or the conjugate map $\lambda \mapsto \overline{\lambda}$.

Also as $\|[X_{32}, X_{22}]\| = \|X_{32}\| = \|[X_{32}, X_{33}]\|$,

$$\begin{aligned} \|\mu_{X_{32}} \mu_{X_{22}} S(X_{32})_\sigma S^{-1}\| &= \|[\phi(X_{32}), \phi(X_{22})]\| \\ &= \|[\phi(X_{32}), \phi(X_{33})]\| = \|\mu_{X_{32}} \mu_{X_{33}} S(X_{32})_\sigma S^{-1}\|. \end{aligned}$$

Thus, $\mu_{X_{22}} = \mu_{X_{33}}$ and from (2.6) and the fact that $\|X_{21}\| = \|X_{31}\|$, we have

$$\|S(X_{21})_\sigma S^{-1}\| = \|S(X_{31})_\sigma S^{-1}\|.$$

We now claim that S is a multiple of some unitary matrix. If not, then there is a pair of orthonormal vectors y_2, y_3 such that $\|Sy_2\| \neq \|Sy_3\|$. Extend $\{y_2, y_3\}$ to an orthonormal basis $\{y_1, y_2, y_3, \dots, y_n\}$ and let $x_j = (y_j)_{\sigma^{-1}}$. Then $\{x_1, \dots, x_n\}$ also forms an orthonormal basis. By the above study, we have

$$\|Sy_2\|\|y_1^* S^{-1}\|\|E_{12}\| = \|S(X_{21})_\sigma S^{-1}\| = \|S(X_{31})_\sigma S^{-1}\| = \|Sy_3\|\|y_1^* S^{-1}\|\|E_{12}\|,$$

which contradicts that $\|Sy_2\| \neq \|Sy_3\|$. Thus, S is a multiple of some unitary matrix. By absorbing the constant term, we may assume that S is unitary. Now for any rank one matrices R and T,

$$\|[R,T]\| = \|[\phi(R), \phi(T)]\| = \|\mu_R \mu_T [R_\sigma, T_\sigma]\| = \mu_R \mu_T \|[R_\sigma, T_\sigma]\|.$$

Since the norm is unitary similarity invariant, $\|[R,T]\| = \|[R_\sigma, T_\sigma]\|$ whenever $[R,T]$ is a rank one nilpotent matrix, and hence $\mu_R \mu_T = 1$ in this case.

Now for any rank one matrix A, we can always find two other rank one matrices B and C such that $[A,B]$, $[A,C]$ and $[B,C]$ are all rank one nilpotents. Then we must have $\mu_A \mu_B = \mu_A \mu_C = \mu_B \mu_C = 1$. As all μ_A, μ_B, μ_C are positive real numbers, the equality is possible only when $\mu_A = \mu_B = \mu_C = 1$. Then we have $\phi(A) = SA_\sigma S^{-1} = SA_\sigma S^*$ for all rank one A.

By replacing ϕ with the map $A \mapsto S^* \phi(A) S$, we may assume that $\phi(X) = X^+$ for all rank one matrices X, where $X^+ = X$, $\overline{X}$, X^t or X^*. Then

$$\|[A,B]\| = \|[\phi(A), \phi(B)]\| = \|[A^+, B^+]\| = \|[A,B]^+\| \quad \text{for all rank one } A, B \in M_n.$$

Notice that the set $\{X : X = [A,B] \text{ for some rank one } A \text{ and } B\}$ contains the set of trace zero non-nilpotent matrices with rank at most two and so is dense in the set of trace zero matrices with rank at most two. Thus, we see that

$$\|X\| = \|X^+\| \quad \text{for all trace zero matrices } X \text{ with rank at most two.}$$

Now define $\Phi : M_n \to M_n$ by $A \mapsto \phi(A)^+$. Then $\Phi(X) = X$ for all rank one matrices X. For any $A \in M_n$ and rank one matrix $X \in M_n$, as $[A,X]$ is a trace zero matrix with rank at most two,

$$\|[A,X]\| = \|[\phi(A), \phi(X)]\| = \|[\phi(A), X^+]\| = \|[\phi(A)^+, X]\| = \|[\Phi(A), X]\|.$$

Thus, $\|[A,X]\| = \|[\Phi(A), X]\|$ for all rank one X. Then Corollary 2.5 implies that $\Phi(A) = \mu_A A + \nu_A I_n$ or $\Phi(A) = \mu_A A^* + \nu_A I_n$ for all diagonalizable matrices A and the latter case happens only when A is normal.

After absorbing the constants μ_A and ν_A, we may assume that $\Phi(X) = X$ for all non-normal diagonalizable matrices X. Then

$$\|[A,B]\| = \|[\phi(A), \phi(B)]\| = \|[\Phi(A), \Phi(B)]^+\| = \|[A,B]^+\|$$

for all non-normal diagonalizable matrices A and B. Since the set of all non-normal diagonalizable matrices is dense in M_n, we see that $\|[A,B]\| = \|[A,B]^+\|$ for all $A, B \in M_n$. Then for any $A \in M_n$,

$$\|[A,X]\| = \|[\phi(A), \phi(X)]\| = \|[\Phi(A), \Phi(X)]^+\| = \|[\Phi(A), \Phi(X)]\| = \|[\Phi(A), X]\|$$

for all non-normal diagonalizable matrices X, and so $\|[A,X]\| = \|[\Phi(A), X]\|$ for all $X \in M_n$. Now the result follows by Lemma 2.6.

Finally, we show that **one only needs the surjective assumption** to get the conclusion on ϕ.

For any $A, B \in M_n$, we say $A \sim B$ if

$$\|[A, X]\| = \|[B, X]\| \quad \text{for all } X \in M_n.$$

Then $\sim$ is an equivalence relation. For each $A \in M_n$, let $S_A = \{B : B \sim A\}$ be the equivalence class of A. By Lemma 2.6, either

(I) S_A is the set of matrices of the form $\mu A + \nu I$ for some $\mu, \nu \in \mathbb{C}$ with $|\mu| = 1$, or

(II) A is normal and $A \sim A^*$, S_A is the set of matrices of the form $\mu A + \nu I$ or $\mu A^* + \nu I$ for some $\mu, \nu \in \mathbb{C}$ with $|\mu| = 1$.

Pick a representative for each equivalence class and write $\mathcal{A}$ for the set of these representatives. Since ϕ is surjective, S_A and $\phi^{-1}(S_A)$ have the same cardinality c for every $A \in \mathcal{A}$. Thus there exists a map $\psi : M_n \to M_n$ which maps $\phi^{-1}(S_A)$ bijectively onto S_A for each $A \in \mathcal{A}$. Clearly ψ is bijective and $\psi(A) \sim \phi(A)$ for all $A \in M_n$. Then, for any $A, B \in M_n$,

$$\|[A, B]\| = \|[\phi(A), \phi(B)]\| = \|[\psi(A), \phi(B)]\| = \|[\psi(A), \psi(B)]\|.$$

That is, ψ is a bijective map satisfying (2.2). By the previous part of our proof of the theorem under the **bijective assumption**, we see that ψ has the desired form. Hence so does ϕ, as $\psi(A) \sim \phi(A)$ implies $\phi(A) = \mu\psi(A) + \nu I$ or $\phi(A) = \mu\psi(A)^* + \nu I$ when $\psi(A)^*$ is normal and $\psi(A)^* \sim \psi(A)$. □

Remark 2.7. We point out that the triangle inequality of the unitary similarity invariant norm has not been used in any part of the proofs in this section. So the result actually holds true for more general unitary similarity invariant functions.

3. Unitarily invariant norm

Using Theorem 2.1, one can give complete descriptions of maps ϕ satisfying $\|[\phi(A), \phi(B)]\| = \|[A, B]\|$ for a specific norm by characterizing the elements in the set $\mathcal{N}_n$ in the theorem. For example, let $\|A\|_F = (\operatorname{tr} A^*A)^{1/2}$ be the Frobenius norm on M_n. We have the following.

Proposition 3.1. *If $N \in M_n$ is normal, then $\|[N, X]\|_F = \|[N^*, X]\|_F$ for all $X \in M_n$. Consequently, if $\|\cdot\|$ is a multiple of the Frobenius norm in Theorem 2.1, then ϕ has the form described there and $\mathcal{N}_n$ can be any subset of the normal matrices.*

Proof. Write $N = U^*DU$ where $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ is diagonal and U is unitary. Notice that $[D, X] = \Lambda \circ X$ where Λ is the skew-symmetric matrix with (i, j)-entry $\lambda_i - \lambda_j$ and $\circ$ denotes the Schur product. Then

$$\begin{aligned} \|[N, X]\|_F &= \|[D, UXU^*]\|_F = \|\Lambda \circ UXU^*\|_F = \|\overline{\Lambda} \circ UXU^*\|_F \\ &= \|[D^*, UXU^*]\|_F = \|[U^*D^*U, X]\|_F = \|[N^*, X]\|_F. \end{aligned}$$

The second assertion is clear. □

Similarly, one can use Theorem 2.1 to characterize the set $\mathcal{N}_n$ for specific norms such as the Schatten p-norm and Ky Fan k-norm on M_n defined by

$$\|A\|_p = \left(\sum_{j=1}^{n} s_j(A)^p\right)^{\frac{1}{p}} \quad \text{and} \quad \|A\|_k = \sum_{j=1}^{k} s_j(A),$$

respectively. Instead of doing a case-by-case study, we prove a general result concerning the characterization of $\mathcal{N}_n$ for unitarily invariant norms. In the following, we always assume that $\|\cdot\|$ is a unitarily invariant norm on M_n not equal to a multiple of the Frobenius norm. We shall always normalize $\|\cdot\|$ so that we may assume $\|E_{11}\| = 1$.

Proposition 3.2. *Let $\|\cdot\|$ be a unitarily invariant norm on M_n, and let*

$$m = \max\{r : \|A\| = \|A\|_F \text{ for all } A \text{ with rank at most } r\}.$$

Suppose $N \in M_n$ is a normal matrix. Then

$$\|[N, X]\| = \|[N^*, X]\| \quad \text{for all } X \in M_n \tag{3.1}$$

if and only if one of the following holds.

(i) *N has collinear or concyclic eigenvalues;*
(ii) *the maximum multiplicity of eigenvalues of N is at least $n - m/2$;*
(iii) *the maximum multiplicity of eigenvalues of N equals $n-(m+1)/2$, provided that for any $A, B \in M_n$ with rank at most $m + 1$, $\|A\| = \|B\|$ whenever*

$$\|A\|_F = \|B\|_F \quad \text{and} \quad \prod_{j=1}^{m+1} s_j(A) = \prod_{j=1}^{m+1} s_j(B).$$

In all other cases, there exists $X \in M_n$ with real distinct eigenvalues such that $\|[N, X]\| \neq \|[N^*, X]\|$.

By the above proposition, we can say more about Theorem 2.1 if the underlying norm is unitarily invariant.

Theorem 3.3. *Suppose $n \geq 3$ and $\phi : M_n \to M_n$ is a surjective map satisfying*

$$\|[A, B]\| = \|[\phi(A), \phi(B)]\|.$$

Then ϕ has the form described in Theorem 2.1 and a normal matrix N is in $\mathcal{N}_n$ only if N satisfies Proposition 3.2(i), (ii) or (iii).

It is clear that for the Schatten p-norm on M_n with $p \neq 2$, $A \in M_n$ satisfies $\|A\|_F = \|A\|_p$ if and only if $\operatorname{rank}(A) \leq 1$. Similarly, for the Ky Fan k-norm on M_n, $A \in M_n$ satisfies $\|A\|_F = \|A\|_k$ if and only if $\operatorname{rank}(A) \leq 1$. Consequently, we have the following.

Corollary 3.4. *If the norm under consideration is the Schatten p-norm with $p \neq 2$ or Ky Fan k-norm, the set $\mathcal{N}_n$ in Theorem 3.3 is a subset of normal matrices with collinear or concyclic eigenvalues only.*

To prove Proposition 3.2, we need the following result.

Theorem 3.5. *[Frobenius-König] Let A be an $n \times n$ matrix. Every diagonal of A contains a zero entry if and only if A has an $r \times s$ zero submatrix such that $r + s > n$.*

Lemma 3.6. *Let $N \in M_n$ be normal with collinear or concyclic eigenvalues. Then $\|[N, X]\| = \|[N^*, X]\|$ for all $X \in M_n$.*

Proof. If N is normal with collinear eigenvalues, then $N^* = \alpha N + \beta I_n$ for some $\alpha, \beta \in \mathbb{C}$ with $|\alpha| = 1$. Then the result clearly follows.

If N is normal with concyclic eigenvalues, $N = \alpha U + \beta I_n$ for some $\alpha, \beta \in \mathbb{C}$ and unitary $U \in M_n$. Then

$$\begin{aligned} \|[N, X]\| = \|[\alpha U, X]\| &= |\alpha| \, \|UX - XU\| = |\alpha| \, \|U^*(UX - XU)U^*\| \\ &= |\overline{\alpha}| \, \|XU^* - U^*X\| = \|[\overline{\alpha}U^*, X]\| = \|[N^*, X]\|. \end{aligned}$$

□

Proof of Proposition 3.2. If N is normal with collinear or concyclic eigenvalues, the result follows by Lemma 3.6.

Now assume that N has neither collinear nor concyclic eigenvalues (so N has at least 4 distinct eigenvalues and $n \geq 4$). Let λ be an eigenvalue of N with maximal multiplicity; write $n - k$ for its multiplicity. Without loss of generality we may replace N by $N - \lambda I$ and assume $\operatorname{rank} N = k$. Note $k \geq 3$.

Case 1. Suppose $2k \leq m$. Since both $[N, X]$ and $[N^*, X]$ have rank at most $2k$, the norm of both is equal to their Frobenius norm. The result follows from Proposition 3.1.

Now suppose $2k > m$. Write $N = UDU^*$ where

$$D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k, 0, \ldots, 0)$$

and all λ_j are nonzero. Without loss of generality, $\lambda_1, \lambda_2, \lambda_3, 0$ are neither collinear nor concyclic. Let Λ be the matrix whose (i, j)-entry is $\lambda_i - \lambda_j$. A *transversal* of an $n \times n$ matrix is defined to be a set of n positions such that no two positions lie in the same row or column.

Case 2. Suppose $2k > m + 1$. It suffices to show that $\|\Lambda \circ X\| \neq \|\overline{\Lambda} \circ X\|$ for some X with distinct real eigenvalues. By way of contradiction, suppose not.

Let Y be a rank $m + 1$ matrix such that $\|Y\| \neq \|Y\|_F$. Let $s_1 \geq \cdots \geq s_{m+1}$ be the nonzero singular values of Y. Let $t = \sqrt{s_1^2 + s_2^2}$.

Define a matrix $X = (X_{ij})$ as follows. Set

$$X_{13} = a/(\lambda_1 - \lambda_3), \quad X_{23} = c/(\lambda_2 - \lambda_3), \quad X_{1,k+1} = b/\lambda_1, \quad X_{2,k+1} = d/\lambda_2$$

(a, b, c, d are free parameters).

Subcase a. If $2k \le n$, define

$$X_{j,k+j-1} = s_j/\lambda_j \quad \text{for } 3 \le j \le k,$$
$$X_{k+j,j} = -s_{k+j}/\lambda_j \quad \text{for } j = 1, 2,$$
$$X_{k+j,j} = -s_{k+j-1}/\lambda_j \quad \text{for } 3 < j \le k.$$

Set all other off-diagonal entries of X equal to zero. Finally, choose values for the diagonal so that X has distinct real eigenvalues (this can be done by [4]).

Subcase b. If $2k > n$, find a transversal of the $(n-2) \times (n-2)$ submatrix obtained by deleting the 1st and 2nd rows, and 3rd and $(k+1)$th columns, of Λ, which avoids any positions in which Λ has a zero. Such a transversal exists by the Frobenius-König theorem (Theorem 3.5). Indeed, the largest zero submatrix of Λ has size $(n-k) \times (n-k)$, so such a transversal exists if $(n-k)+(n-k) \le n-2$, or $2k \ge n+2$. If $2k = n+1$, there is only one $(n-k) \times (n-k)$ zero submatrix of Λ (since N has at least 4 distinct eigenvalues). Since we have deleted the $(k+1)$th column of X, the largest forbidden submatrix has size $(n-k) \times (n-k-1)$. Since $(n-k)+(n-k-1) \le n-2$ if and only if $2k \ge n+1$, we can again find such a transversal.

Define the entries of X on this transversal to be the singular values $s_3, \ldots, s_n$ of Y divided by the entries of Λ in the corresponding positions. Set all other off-diagonal entries of X equal to zero. Choose values for the diagonal so that X has distinct real eigenvalues.

In either case, the singular values of $\Lambda \circ X$ (respectively $\overline{\Lambda} \circ X$) are given by the singular values $s_3, \ldots, s_n$ of Y, together with the singular values of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (respectively $\begin{bmatrix} ae^{-i\alpha} & be^{-i\beta} \\ ce^{-i\gamma} & de^{-i\delta} \end{bmatrix}$, where $\alpha = 2\arg(\lambda_1 - \lambda_3)$, $\beta = 2\arg\lambda_1$, $\gamma = 2\arg(\lambda_2 - \lambda_3)$, and $\delta = 2\arg\lambda_2$. Given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $z \in \mathbb{C}$, write

$$F(A, z) = \left\| \begin{bmatrix} a & b \\ c & dz \end{bmatrix} \oplus \operatorname{diag}(s_3, \ldots, s_n) \right\|.$$

Letting $\theta = -\alpha - \delta + \beta + \gamma$, we have $F(A, 1) = F(A, e^{i\theta})$ for any $A \in M_2$ (by our initial assumption $\|\Lambda \circ X\| = \|\overline{\Lambda} \circ X\|$). Replacing d by $de^{i\theta}$ we see that $F(A, 1) = F(A, e^{ik\theta})$ for all $k \in \mathbb{Z}$.

Note that

$$A_\phi = \frac{t}{\sqrt{2}} \begin{bmatrix} \cos\phi & \cos\phi \\ \sin\phi & \sin\phi \end{bmatrix}$$

has singular values t and 0, so $F(A_\phi, 1) = \|Y\|_F$. Since any pair (σ_1, σ_2) satisfying $\sigma_1^2 + \sigma_2^2 = t^2$ can be the singular values of $A_{\pi/4} \circ \begin{bmatrix} 1 & 1 \\ 1 & z \end{bmatrix}$ for some complex unit z, we are done if θ/π is irrational (by continuity, we can make $\|Y\|_F = F(A_{\pi/4}, e^{ik\theta})$

arbitrarily close to $\|Y\|$, giving a contradiction). Otherwise choose $k_0 \in \mathbb{Z}$ so that $\omega = e^{ik_0\theta}$ is as close to -1 as possible. (Since $0, \lambda_1, \lambda_2, \lambda_3$ are neither collinear nor concyclic, θ is not a multiple of 2π.) We have $\arg\omega \in [2\pi/3, 4\pi/3]$, so $|1+\omega| \le 1$.

Since $\|Y\| = F(A_\phi, 1) = F(A_\phi, \omega)$ for any $\phi \in [0, 2\pi]$, the norm is constant on all matrices with singular values equal to

$$p, q, s_3, \ldots, s_n$$

where $p^2 + q^2 = t^2$, $p \ge q$, and $p \ge t\sqrt{3}/2$, $q \le t/2$. Writing $B = \begin{bmatrix} t/2 & t/2 \\ t/2 & t/2 \end{bmatrix}$, it follows that

$$\|Y\|_F = F(B, e^{i\psi}) = F(B, e^{i\psi}e^{ik\theta})$$

for all $0 \le \psi \le 2\pi/3$ and $k \in \mathbb{Z}$. Choose ψ, k so that the singular values of $B \circ \begin{bmatrix} 1 & 1 \\ 1 & e^{i(\psi+k\theta)} \end{bmatrix}$ are s_1, s_2. Then $\|Y\|_F = \|Y\|$, a contradiction.

Case 3. Now suppose $2k = m+1$. We divide into two subcases.

Subcase a. Suppose $\|A\| = \|B\|$ whenever

$$\operatorname{rank} A,\ \operatorname{rank} B \le 2k, \quad \|A\|_F = \|B\|_F, \text{ and } \quad \prod_{j=1}^{2k} s_j(A) = \prod_{j=1}^{2k} s_j(B).$$

Recall that N has rank k. Then $\|[N, X]\| = \|[N^*, X]\|$ for all $X \in M_n$ since $[N, X]$ and $[N^*, X]$ both have rank at most $2k$, have the same Frobenius norm, and the products of the largest $2k$ singular values are the same. To see this last assertion, we may assume $N = D \oplus 0_{n-k}$, where $D \in M_k$ is a nonsingular diagonal matrix, and write $X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$ where $X_{11} \in M_k$. Using QR decompositions, we may write $X_{12} = \begin{bmatrix} B & 0 \end{bmatrix} V$ and $X_{21} = U \begin{bmatrix} C \\ 0 \end{bmatrix}$ where $U, V \in M_{n-k}$ are unitary and $B, C \in M_k$. Then

$$[N, X] = \begin{bmatrix} [D, X_{11}] & DX_{12} \\ -X_{21}D & 0 \end{bmatrix} = \begin{bmatrix} I_k & 0 \\ 0 & U \end{bmatrix} \begin{bmatrix} [D, X_{11}] & D\begin{bmatrix} B & 0 \end{bmatrix} \\ -\begin{bmatrix} C \\ 0 \end{bmatrix} D & 0_{n-k} \end{bmatrix} \begin{bmatrix} I_k & 0 \\ 0 & V \end{bmatrix},$$

so

$$\prod_{j=1}^{2k} s_j([N, X]) = |(\det DB)(\det CD)| = |(\det D^*B)(\det CD^*)| = \prod_{j=1}^{2k} s_j([N^*, X])$$

as claimed.

Subcase b. Suppose Case 3a does not holds. That is, there exist positive numbers $a_1, \ldots, a_{2k}$ and $b_1, \ldots, b_{2k}$ such that

$$\sum_{j=1}^{2k} a_j^2 = \sum_{j=1}^{2k} b_j^2 \quad \text{and} \quad \prod_{j=1}^{2k} a_j = \prod_{j=1}^{2k} b_j$$

but $\|\mathrm{diag}\,(a_1,\ldots,a_{2k})\| \neq \|\mathrm{diag}\,(b_1,\ldots,b_{2k})\|$. Without loss of generality we may assume $a_j = b_j$ for $j \geq 4$.

Let $C = a_1^2 + a_2^2 + a_3^2$ and $D = a_1^2 a_2^2 a_3^2$, and define

$$\Omega = \{(x,y,z) : x^2+y^2+z^2 = C,\ x^2y^2z^2 = D, \text{ and } x \geq y \geq z > 0\}.$$

Let $s_{min} = \min\{x : (x,y,z) \in \Omega \text{ for some } y,z\}$ and $s_{max} = \max\{x : (x,y,z) \in \Omega \text{ for some } y,z\}$. One sees that $\Omega = \{(x, y(x), z(x)) : x \in [s_{min}, s_{max}]\}$ where $y(x), z(x)$ are the unique (continuous) functions such that $(x, y(x), z(x)) \in \Omega$. Let $h(x) = \|\mathrm{diag}\,(x, y(x), z(x), a_4, \ldots, a_{2k})\|$.

Since h is continuous but not constant on the interval $[s_{min}, s_{max}]$, there is some $\tau \in (s_{min}, s_{max})$ such that $h(\tau) \neq h(s_{min})$. Let

$$t = \inf\{x \in [s_{min}, s_{max}] : h(x) = h(\tau)\}.$$

Note that $s_{min} < t$, since, by continuity of h, $h(t) = h(\tau)$.

Clearly, the set

$$\left\{ s_1(Y) : \|Y\|_F^2 = C, |\det Y|^2 = D, Y = \begin{bmatrix} 0 & 0 & c \\ 0 & f & 0 \\ g & 0 & 0 \end{bmatrix},\ c,f,g > 0 \right\}$$

is just $[s_{min}, s_{max}]$. By continuity, (s_{min}, s_{max}) is a subset of

$$\left\{ s_1(A) : \|A\|_F^2 = C,\ |\det A|^2 = D,\ A = \begin{bmatrix} a & b & c \\ d & f & 0 \\ g & 0 & 0 \end{bmatrix},\ a,b,c,d,f,g > 0 \right\},$$

so we can find $A = \begin{bmatrix} a & b & c \\ d & f & 0 \\ g & 0 & 0 \end{bmatrix}$ with $a,\ldots,g > 0$, $\|A\|_F^2 = C$, $|\det A|^2 = D$, and $s_1(A) = t$.

Define a matrix $X = (X_{ij})$ as follows. Set

$X_{13} = a/(\lambda_1 - \lambda_3), \quad X_{1,k+1} = b/\lambda_1, \quad X_{1,k+2} = c/\lambda_1,$

$X_{23} = d/(\lambda_2 - \lambda_3), \quad X_{2,k+1} = f/\lambda_2, \quad X_{k+1,3} = -g/\lambda_3,$

$X_{j,k+j} = a_{j+1}/\lambda_j \quad \text{for } 3 \leq j \leq k,$

$X_{k+2,1} = -a_{k+2}/\lambda_1, \quad X_{k+3,2} = -a_{k+3}/\lambda_2, \quad X_{k+j,j} = -a_{k+j}/\lambda_j \quad \text{for } 4 \leq j \leq k.$

Set all other off-diagonal entries of X equal to zero. Choose values for the diagonal so that X has distinct real eigenvalues. Then the singular values of $\Lambda \circ X$ (respectively $\overline{\Lambda} \circ X$) are given by $a_4, \ldots, a_n$, together with the singular values of A (respectively $B = \begin{bmatrix} a & b & c \\ d & fe^{i\theta} & 0 \\ g & 0 & 0 \end{bmatrix}$ where θ is as defined in Case 2b, so $e^{i\theta} \neq 1$).

By way of contradiction, suppose $\|\Lambda \circ X\| = \|\overline{\Lambda} \circ X\|$. We have

$$h(t) = \|\Lambda \circ X\| = \|\overline{\Lambda} \circ X\| = h(s_1(B)),$$

so $s_1(B) \geq t$ by the definition of t. On the other hand, we have

$$s_1(B)^2 = \rho(B^*B) < \rho(A^*A) = s_1(A)^2 = t^2$$

since $|B^*B| \leq A^*A$ but $|B^*B| \neq A^*A$ (see the proof of Theorem 8.4.5 in [6]; here $|C|$ denotes the matrix with (i,j)-entry $|C_{ij}|$), giving the desired contradiction. $\square$

4. k-NUMERICAL RADIUS

If the norm in Theorem 2.1 is unitary similarity invariant but not unitarily invariant, it is not so easy to characterize the set $\mathcal{N}_n$. In the following, we consider a class of unitary similarity invariant norms and show that the set $\mathcal{N}_n$ in Theorem 2.1 has to be empty. Recall that for $1 \leq k < n$, the *k-numerical range* of $A \in M_n$ is the set

$$W_k(A) = \{\mathrm{tr}\,(AP) : P \in M_n, P^2 = P = P^*,\ \mathrm{tr}\,P = k\},$$

and the *k-numerical radius* of A is the quantity

$$w_k(A) = \max\{|\mu| : \mu \in W_k(A)\}.$$

Notice that $w_k(\cdot)$ is a unitary similarity invariant norm but not a unitarily invariant norm. We have the following result.

Theorem 4.1. *Suppose $1 \leq k < n$, and $\phi : M_n \to M_n$ is a surjective map satisfying*

$$w_k([\phi(A), \phi(B)]) = w_k([A,B]) \quad \text{for all } A, B \in M_n.$$

Then ϕ has the form described in Theorem 2.1 with $\mathcal{N}_n = \emptyset$. That is,

$$\phi(A) = \mu_A U A^\dagger U^* + \nu_A I \text{ for all } A \in M_n$$

for some unitary U, $\mu_A, \nu_A \in \mathbb{C}$ with $|\mu_A| = 1$, depending on A, and $A^\dagger = A$, $\overline{A}$, A^t or A^*.

By Theorem 2.1, we only need to prove that $\mathcal{N}_n$ has to be empty. We start with the following lemma.

Lemma 4.2. *Suppose $C \in M_3$ has trace zero and $\mathrm{tr}\,(C^*C) = 1$. Then*

$$w_1(C) = w_2(C) \leq \sqrt{2/3}.$$

The equality holds if and only if C is unitarily similar to $\xi(J_3 - I_3)$ for some $\xi \in \mathbb{C}$ with $|\xi| = 1/\sqrt{6}$. (Recall J_n is the $n \times n$ matrix whose every entry is one.)

Proof. Since $\mathrm{tr}\,C = 0$, we have $\mu \in W_1(C)$ if and only if $-\mu \in W_2(C)$. So, $w_1(C) = w_2(C)$.

For each $t \in [0, 2\pi)$, let $H_t = (e^{it}C + e^{-it}C^*)/2$. Then $\mathrm{tr}\,(H_t) = 0$ and $\mathrm{tr}\,(H_t^2) \leq \mathrm{tr}\,(C^*C) = 1$. Thus, H_t has eigenvalues $h_1 \geq h_2 \geq h_3$ satisfying $h_1 + h_2 + h_3 = 0$ and $h_1^2 + h_2^2 + h_3^2 \leq 1$. It is easy to show that

$$w_1(H_t) = \max\{h_1, -h_3\} \leq \sqrt{2/3},$$

and the equality holds if and only if $(h_1, h_2, h_3) = \pm(2, -1, -1)/\sqrt{6}$. Consequently,

$$w_1(C) = \max\{w_1(H_t) : t \in [0, 2\pi)\} \le \sqrt{2/3},$$

and the equality holds if and only if there is $t \in [0, 2\pi)$ such that H_t has eigenvalues $(2, -1, -1)/\sqrt{6}$, that is, H_t is unitarily similar to $(J_3 - I_3)/\sqrt{6}$. Note that $e^{it}C = H_t + iG_t$ such that $\mathrm{tr}\,(C^*C) = \mathrm{tr}\,(H_t^2) + \mathrm{tr}\,(G_t^2)$. As $\mathrm{tr}\,(C^*C) = \mathrm{tr}\,(H_t^2) = 1$, $\mathrm{tr}\,G_t^2 = 0$ so that $G_t = 0$. Thus, $w_1(C) = \sqrt{2/3}$ if and only if there is $t \in [0, 2\pi)$ such that $e^{it}C = H_t$ is unitarily similar to $(J_3 - I_3)/\sqrt{6}$. □

Proof of Theorem 4.1. First, suppose $\mathcal{N}_n$ contains a normal matrix with collinear eigenvalues N. Then $N^* = \alpha N + \beta I$ for some $\alpha, \beta \in \mathbb{C}$ with $|\alpha| = 1$, and so $\phi(N) = \mu_N(N^\dagger)^* + \nu_N I = \hat{\mu}_N(N^\dagger) + \hat{\nu}_N I$ where $\hat{\mu}_N = \alpha\mu_N$ and $\hat{\nu}_N = \beta\mu_N + \nu_N$ if $N^\dagger = N$ or N^t, or $\hat{\mu}_N = \overline{\alpha}\mu_N$ and $\hat{\nu}_N = \overline{\beta}\mu_N + \nu_N$ if $N^\dagger = \overline{N}$ or N^*. In all cases, we can replace $\mathcal{N}_n$ by $\mathcal{N}_n \setminus \{N\}$. Thus, we may assume that $\mathcal{N}_n$ does not contain any normal matrix with collinear eigenvalues.

Now assume $\mathcal{N}_n$ contains a normal matrix A with at least three non-collinear eigenvalues. Applying a unitary similarity, we may assume that $A = \mathrm{diag}\,(a_1, \dots, a_n)$ where a_1, a_2, a_3 are the three distinct non-collinear points.

For any non-normal $X \in M_n$, $w_k([A, X]) = w_k([\phi(A), \phi(X)]) = w_k([A^*, X])$. Since this is true for all non-normal $X \in M_n$, by continuity of the map $X \mapsto w_k(X)$ and the fact that the set of non-normal matrices is dense in M_n, we see that $w_k([A, X]) = w_k([A^*, X])$ for any $X \in M_n$.

Let

$$B = \sqrt{1/6}\begin{bmatrix} 0 & 1/(a_1 - a_2) & 1/(a_1 - a_3) \\ 1/(a_2 - a_1) & 0 & 1/(a_2 - a_3) \\ 1/(a_3 - a_1) & 1/(a_3 - a_2) & 0 \end{bmatrix} \oplus 0_{n-3}.$$

Then $[A, B] = (J_3 - I_3)/\sqrt{6} \oplus 0_{n-3}$ has eigenvalues $2/\sqrt{6}, -1/\sqrt{6}, -1/\sqrt{6}, 0, \dots, 0$. Note that if H is Hermitian with eigenvalues $h_1 \ge \cdots \ge h_n$, then

$$w_k(H) = \max\left\{\sum_{j=1}^{k} h_j, -\sum_{j=1}^{k} h_{n-j+1}\right\}. \tag{4.1}$$

Hence $w_k([A, B]) = \sqrt{2/3}$. Now, $[A^*, B] = C \oplus 0_{n-3}$ with

$$C = \sqrt{1/6}\begin{bmatrix} 0 & b & c \\ b & 0 & d \\ c & d & 0 \end{bmatrix},$$

where $b = (\bar{a}_1 - \bar{a}_2)/(a_1 - a_2)$, $c = (\bar{a}_1 - \bar{a}_3)/(a_1 - a_3)$ and $d = (\bar{a}_2 - \bar{a}_3)/(a_2 - a_3)$ are complex units. If $k = 1$, then $w_1(C \oplus 0_{n-3}) = w_1(C)$. If $1 < k < n$, then

$$w_k(C \oplus 0_{n-3}) = \max\left\{w_k\left((e^{it}C + e^{-it}C^*) \oplus 0_{n-3}\right)/2 : t \in [0, 2\pi)\right\} = w_2(C).$$

Note that C satisfies the hypothesis of Lemma 4.2, and $w_1(C) = w_2(C) = \sqrt{2/3}$. Thus, there is $\mu \in \mathbb{C}$ with $|\mu| = 1$ such that μC is Hermitian with eigenvalues

$(2,-1,-1)/\sqrt{6}$. Replacing (A,B) by $(A,B)/\sqrt{\mu}$, we may assume that $\mu = 1$. So, C is Hermitian, and we have $b^2 = c^2 = d^2 = 1$. Thus, $b, c, d \in \{1, -1\}$, and two of the real values in $\{b, c, d\}$ are equal. Without loss of generality, assume $b = c$.

Case 1. If $b = c = 1$, then $\bar{a}_1 - \bar{a}_2 = a_1 - a_2$ and $\bar{a}_1 - \bar{a}_3 = a_1 - a_3$. So, both $a_1 - a_2$ and $a_1 - a_3$ are real. It follows that a_1, a_2, a_3 are collinear. This contradicts the fact that a_1, a_2, a_3 are non-colinear.

Case 2. If $b = c = -1$, then $\bar{a}_1 - \bar{a}_2 = -(a_1 - a_2)$ and $\bar{a}_1 - \bar{a}_3 = -(a_1 - a_3)$. Thus, $a_1 - a_2$ and $a_1 - a_3$ are real multiples of i. It follows that a_1, a_2, a_3 are collinear. Contradiction arrived.

So, we see that $\mathcal{N}_n$ cannot contain a matrix with three non-collinear eigenvalues and so $\mathcal{N}_n$ is empty. □

References

1. J.T. Chan, C.K. Li, and N.S. Sze, Isometries for unitarily invariant norms, Linear Algebra Appl. 399 (2005), 53-70.
2. J.T. Chan, C.K. Li, and N.S. Sze, Mappings on matrices: Invariance of functional values of matrix products, J. of Aust. Math. Soc., 81 (2006), 165-184.
3. W.S. Cheung, S. Fallat and C.K. Li, Multiplicative preservers on semigroups of matrices, Liner Algebra Appl. 355 (2002), 173-186.
4. S. Friedland, Matrices with prescribed off-diagonal elements. Israel J. Math. 11 (1972), 184-189.
5. R. Guralnick and C.K. Li, Invertible preservers and algebraic groups III: preservers of unitary similarity (congruence) invariants and overgroups of some unitary subgroups, Linear and Multilinear Algebra 43 (1997), 257-282.
6. R. Horn and C. Johnson, Matrix analysis. Cambridge University Press, Cambridge, 1985.
7. P. Šemrl, Non-linear commutativity preserving maps. Acta Sci. Math. (Szeged) 71 (2005), 781–819.

Department of Mathematics, College of William & Mary, Williamsburg, VA 23185
E-mail address: `ckli@math.wm.edu`

Department of Mathematics, Embry-Riddle Aeronautical University, Prescott, AZ 86301
E-mail address: `poon3de@erau.edu`

Department of Mathematics, University of Connecticut, Storrs, CT 06269
E-mail address: `sze@math.uconn.edu`